

Claims

I claim:

1. A method for selecting rows from a partitioned database table, comprising the steps of:
 - a. creating a file context, which stores at least location data for a row and a first value associated with the row, for each populated partition of the partitioned database table;
 - b. determining the lowest first value stored by the file contexts; and
 - c. identifying rows with a particular first value by reading the file contexts.
2. The method of claim 1, where rows are ordered within the populated partitions in accordance with the first value for each row.
3. The method of claim 1, where the first values stored are based at least in part on values in one or more specified columns of the associated rows.
4. The method of claim 3, where the first values stored by the file contexts are the result of a hash function applied to values in one or more specified columns of the associated rows.
5. The method of claim 1, where step a comprises the steps of:
 - i. creating an empty file context for a populated partition;
 - ii. finding the lowest first value associated with a row in that partition;
 - iii. storing the location data and the first value for the row having the lowest first value in the file context for that partition; and
 - iv. repeating steps i – iii until file contexts have been created for all populated partitions.

6. The method of claim 1, where step c comprises the steps of:
- i. examining file contexts in order of the populated partitions until a file context storing the particular first value is selected;
 - ii. reading the location data for the row;
 - 5 iii. reading the row data;
 - iv. if there is at least one additional row in the partition, comparing the first value for the next row in the partition to the particular first value;
 - v. repeating steps ii – iv if the first value for the next row equals the particular first value; and
 - 10 vi. repeating steps i – v until all file contexts for the table have been examined.
7. The method of claim 1, further comprising the step of:
- d. updating each file context to store location data and the first value for a row in the associated partition with a first value equal to or greater than a desired first value.
8. The method of claim 7, where updating each file context comprises:
- i. examining file contexts in order of the populated partitions until a file context storing a first value less than the desired first value is selected;
 - ii. if the partition corresponding to the selected file context contains at least one row having the desired first value, storing the location data and first value for the first such row in the file context;
 - 20 iii. if the partition corresponding to the selected file context contains no row having the desired first value and at least one row having a first value greater than the desired first value, storing the location data and first value for the first such row in the file context;
 - iv. if the partition corresponding to the selected file context contains no row having the desired first value and no row having a first value greater than the desired first value, removing the file context; and
 - 25 v. repeating steps i – iv until all file contexts for the table have been examined.

9. A database system for selecting rows from a partitioned database table, the partitioned database table including rows and columns and being divided into partitions and at least one of the partitions in the table being populated by one or more rows, the system comprising:

one or more nodes;

5 a plurality of CPUs, each of the one or more nodes providing access to one or more CPUs;

a plurality of processes, each of the one or more CPUs providing access to one or more virtual processes;

10 each process configured to manage data, including the partitioned database table, stored in one of a plurality of data-storage facilities;

a partitioned table access component configured to select rows from the table by:

- a. creating a file context, which stores at least location data for a row and a first value associated with the row, for each populated partition;
- b. determining the lowest first value stored by the file contexts; and
- c. identifying rows with a particular first value by reading the file contexts.

10. The system of claim 9, where the rows of the table are stored across at least two of the data-storage facilities and are partitioned within each data-storage facility.

11. The system of claim 9, where rows are ordered within the populated partitions in accordance with the first value for each row.

20 12. The system of claim 9, where the first values stored are based at least in part on values in one or more specified columns of the associated rows.

13. The system of claim 9, where the first values stored by the file contexts are the result of a hash function applied to values in one or more specified columns of the associated rows.

14. The system of claim 9, where step a comprises the steps of:
- i. creating an empty file context for a populated partition;
 - ii. finding the lowest first value associated with a row in that partition;
 - iii. storing the location data and the first value for the row having the lowest first value in
5 the file context for that partition; and
 - iv. repeating steps i – iii until file contexts have been created for all populated partitions.
15. The system of claim 9, where step c comprises the steps of:
- i. examining file contexts in order of the populated partitions until a file context storing
the particular first value is selected;
 - 10 ii. reading the location data for the row;
 - iii. reading the row data;
 - iv. if there is at least one additional row in the partition, comparing the first value for the
next row in the partition to the particular first value;
 - v. repeating steps ii – iv if the first value for the next row equals the particular first
value; and
 - vi. repeating steps i – v until all file contexts for the table have been examined.
16. The system of claim 9, where the partitioned table access component is configured to select
rows from the table by updating each file context to store location data and the first value for a
row in the associated partition with a first value equal to or greater than a desired first value.

17. The system of claim 16, where updating each file context comprises:

- i. examining file contexts in order of the populated partitions until a file context storing a first value less than the desired first value is selected;
- ii. if the partition corresponding to the selected file context contains at least one row having the desired first value, storing the location data and first value for the first such row in the file context;
- iii. if the partition corresponding to the selected file context contains no row having the desired first value and at least one row having a first value greater than the desired first value, storing the location data and first value for the first such row in the file context;
- iv. if the partition corresponding to the selected file context contains no row having the desired first value and no row having a first value greater than the desired first value, removing the file context; and
- v. repeating steps i – iv until all file contexts for the table have been examined.

18. A computer program, stored in a tangible medium, for selecting rows from a partitioned database table, the program comprising executable instructions that cause a computer to:

- a. create a file context, which stores at least location data for a row and a first value associated with the row, for each populated partition of the partitioned database table;
- b. determine the lowest first value stored by the file contexts; and
- c. identify rows with a particular first value by reading the file contexts.

19. The computer program of claim 18, where rows are ordered within the populated partitions in accordance with the first value for each row.

20. The computer program of claim 18, where the first values stored are based at least in part on values in one or more specified columns of the associated rows.

21. The computer program of claim 18, where the first values stored by the file contexts are the result of a hash function applied to values in one or more specified columns of the associated rows.

22. The computer program of claim 18, where step a comprises:

- i. creating an empty file context for a populated partition;
- ii. finding the lowest first value associated with a row in that partition;
- iii. storing the location data and the first value for the row having the lowest first value in the file context for that partition; and
- iv. repeating steps i – iii until file contexts have been created for all populated partitions.

23. The computer program of claim 18, where step c comprises:

- i. examining file contexts in order of the populated partitions until a file context storing the particular first value is selected;
- ii. reading the location data for the row;
- iii. reading the row data;
- iv. if there is at least one additional row in the partition, comparing the first value for the next row in the partition to the particular first value;
- v. repeating steps ii – iv if the first value for the next row equals the particular first value; and
- vi. repeating steps i – v until all file contexts for the table have been examined.

24. The computer program of claim 18, further comprising executable instructions that cause a computer to:

- d. update each file context to store location data and the first value for a row in the associated partition with a first value equal to or greater than a desired first value.

25. The computer program of claim 24, where step d comprises:

- i. examining file contexts in order of the populated partitions until a file context storing a first value less than the desired first value is selected;
- ii. if the partition corresponding to the selected file context contains at least one row having the desired first value, storing the location data and first value for the first such row in the file context;
- iii. if the partition corresponding to the selected file context contains no row having the desired first value and at least one row having a first value greater than the desired first value, storing the location data and first value for the first such row in the file context;
- iv. if the partition corresponding to the selected file context contains no row having the desired first value and no row having a first value greater than the desired first value, removing the file context; and
- v. repeating steps i – iv until all file contexts for the table have been examined.